



Mikkeli Campus

HEDONIC REGRESSION MODEL FOR COFFEE FUTURES

Analysis of effect of weather, exchange rate, past price and spot price

Nguyen Chau Thuy Tien

Bachelor's Thesis

Instructor: Dr. David Volkman

Date of submission: March 20, 2017

Declaration

By completing this cover sheet and declaration, I confirm that this assignment is my own work, is not copied from the work (published or unpublished) of any other person, and has not previously been submitted for assessment either at Aalto University, or another educational establishment. Any direct or indirect uses of material (e.g.: text, visuals, ideas...) from other sources have been fully acknowledged and cited according to the conventions of the Harvard Referencing System.

HEDONIC REGRESSION MODEL FOR COFFEE FUTURES

Analysis of effect of weather, exchange rate, past price and spot price

Nguyen Chau Thuy Tien

Bachelor's Thesis

Instructor: Dr. David Volkman

Date of submission: March 20, 2017

Declaration

By completing this cover sheet and declaration, I confirm that this assignment is my own work, is not copied from the work (published or unpublished) of any other person, and has not previously been submitted for assessment either at Aalto University, or another educational establishment. Any direct or indirect uses of material (e.g.: text, visuals, ideas...) from other sources have been fully acknowledged and cited according to the conventions of the Harvard Referencing System.

COVER PAGE

TITLE PAGE

ABSTRACT

TABLE OF CONTENTS

1. INTRODUCTION

- 1.1. Background
- 1.2. Research Problem
- 1.3. Research Questions
- 1.4. Research Objectives
- 1.5. Importance and significance of the research

2. Definition of key terms

- 2.1. Coffee
- 2.2. Coffee futures
- 2.3. Hedonic price model
- 2.4. Efficient market hypothesis
- 2.5. Price discovery

3. LITERATURE REVIEW

- 3.1. Introduction
- 3.2. Background
- 3.3. Previous studies of coffee futures pricing
- 3.4. Market efficiency
- 3.5. Price discovery
- 3.6. Hedonic regression model
 - 3.6.1 Weather
 - 3.6.2 Exchange rate
 - 3.6.3 Spot price and futures price
 - 3.6.4 Autocorrelation of price
- 3.7. Conclusion

3.8. Conceptual framework

4. METHODOLOGY

4.1. Theoretical model

4.2. Empirical model

4.3. Variable selection

4.3.1 Independent variables

4.3.2 Dummy variable

4.3.3 Dependent variable

4.4. Unit root test

4.5. Suggested empirical coffee futures model

5. FINDINGS

5.1. Unit root test

5.2. Tests for multiple linear regression

5.2.1 Normal distribution test

5.2.2 Linearity test

5.2.3 Correlation tests

5.2.4 Homoscedasticity test

5.3. Descriptive statistics

5.4. Hedonic regression model

6. DISCUSSION AND ANALYSIS

7. CONCLUSIONS

7.1. Main Findings

7.2. Suggestions for Further Research

8. REFERENCES

Author: Nguyen Chau Thuy Tien

Title of thesis:

Date: 13 April 2017

Degree: Bachelor of Science in Economics and Business Administration

Supervisor: Dr. David Volkman

Objectives

The first objective of the thesis is determining the form of time series data that would be suitable for conducting empirical research on coffee futures. The second objective is to conduct a hedonic regression model and to test the observed data for multiple linear regression. Finally, this paper would like to understand the relative importance of economic factors affecting coffee futures price.

Summary

Unit root test is conducted for 228 samples of Coffee "C" Futures and the independent variables: precipitation in Brazil and in Colombia, exchange rates USD/BRL and USD/COD, spot coffee price and coffee futures price of the previous month from January 1994 to December 2012. The Augmented Dicker-Fuller test suggests the first differencing method to remove unit root in time series data.

Several tests for ensuring multiple regression assumptions are performed. The input variables could be said to be valid for conducting hedonic regression model.

Conclusions

The hedonic regression model reveals that 96% of coffee futures price could be explained by precipitation in Brazil and in Colombia, exchange rates USD/BRL and USD/COD, spot coffee price, coffee futures price of the previous month and dummy variable shock. Only spot price is of great significance at 95% confidence, while the others are not. Spot price also has the highest coefficient, while coefficients of other independent variables are relatively low.

Key words: *coffee futures, hedonic regression model, market efficiency*

See: <http://web.lib.aalto.fi/en/helevoc/pdf/>

Language: English

Grade:

1. INTRODUCTION

1.1 Background

Commodity futures is a form of financial markets. Buyers and sellers of commodity futures contracts use them to mitigate the risks associated with the fluctuations of the price of raw material in the future. Besides hedgers and producers, traders, especially speculators could take advantage of price volatility as an opportunity for earning great profits by predicting in which direction (increasing or decreasing) commodity futures will move. Speculating in futures market can involve significant risks due to the high level of leverage when traders would like to hold a large number of contracts.

Commodity futures markets usually react strongly and almost instantly to information and news. Price of commodity futures reflects not only the market conditions but also the change in beliefs and emotions of traders actively involved in the market. As a result, participants (speculators, hedgers, and even regulatory bodies like the government) try to forecast price movements for their own purpose. Hedgers who are often raw material producers, farmers, importers, and exporters avoid price risks by securing a certain price for their products in the futures. Speculators, on the other hand, want to maximize their potential gains through every change in commodity futures price. Both of them are two sides of the market that can benefit each other while the price goes up and down because of the invasion of flooding information. The arbitrage also aims to forecast the price for the purpose of budget planning and resources allocation. Turnovsky (1983) and Choe (1990) both demonstrated that futures price has a close link with commodity prices in the futures. Futures price could be a determinant of the prediction of future spot prices. Especially in the case of primary commodity products, apart from being traded regularly in the exchange market, they are the main source of revenues for exporting countries (Milas, 2000). The contribution of commodity exports to the national gross domestic product (GDP) in developing countries is often greater than developed countries.

The international coffee industry has experienced a steady growth over the two decades (Euromonitor International, 2016: 4). Both production and consumption volume increased due to the appearance of emerging coffee-producing countries and the expansion of coffee drinking culture towards Asia, Middle East, and Africa.

However, key producers such as Brazil, Colombia, and Vietnam continue to evolve. Nevertheless, export-intensive nations tend to undergo remarkable negative economic and financial damage when instability occurs. It can be external shocks such as new policy/act/agreement introduced or crisis happening. This has prompted calls for the sustainable development of the coffee sector and diversification in commodity production.

1.2 Research problem

Whether the financial market is predictable or not has been a quest for numerous quantitative, empirical and even qualitative studies. From Bachelier's financial theory of random walk (1900) to Fama's theory of efficient market (1970) claimed that price is unpredictable. However, a definite number of researchers from different fields have suggested and improved the idea of price modeling, determination and forecasting for product market, stock market, futures market and even options (Black, 1976; Schwartz, 1997; Otero & Milas, 2000; Ajao, 2013).

There is a gap between academic research and practical finance. Most futures pricing methods are heavily mathematical and require advanced knowledge of statistics and quantitative finance or sophisticated software to be computed. This might not be accessible to a large proportion of traders in the futures market. An alternative method of forecasting spurs this research paper.

Traders encounter or pay costs for acquiring a substantial amount of data from both mainstream media channels to exclusive reports from companies, organizations and governmental bodies. The question of which data players in coffee futures market pay more attention to gauge this explanatory thesis. Supply and demand are considered main attributes to the formation and movement of the coffee futures market. Other factors might also contribute to the coffee futures behavior such as weather, exchange rate, spot price, and past price.

1.3 Research questions

Q1: Is it possible to use less fundamental indicators such as weather, exchange rate, spot price, and past price to predict coffee futures price?

Q2: Which type of regression model is appropriate for analyzing the effect of those factors?

1.4 Research objectives

To assess the above questions, a statistical research is conducted. There are three objectives of this study. The first objective is determining the form of time series data that would be suitable for conducting empirical research on coffee futures. The second objective is to conduct a hedonic regression model and to test the observed data for multiple linear regression. Finally, this paper would like to understand the relative importance of economic factors affecting coffee futures price.

1.5 Importance and significance of the research

Finance in general and commodities market in particular is one of the most crucial fields of research, especially to predict the movement of the market (Wolfers & Zitzewitz, 2004). The thesis aims to contribute to the finance sector and international business by tackling with one of the most popular trading approach: fundamental analysis on the underlying forces of market movement.

The futures market has been traded internationally through the availability of Internet trading. Fundamental analysis is globally used because of its high level of information accessibility through Internet data sources. Nevertheless, there have been a limited number of academic studies treating coffee futures as their main subject of empirical research. Therefore, using hedonic regression model to understand the nature and link among coffee futures and its factors might provide a suggestion for further investigation on this topic.

2. DEFINITION

2.1 Coffee

Coffee is considered as one of the top traded commodities globally. Two main types of coffee are Arabica (about 70% of the quantity produced globally) and Robusta (Otero et al., 2000). The two main Arabica producers are Brazil and Colombia which account for more than 50% of world production. The International Trade Centre report named

Bitter or better future for coffee producers (2004) highlights that low production costs thanks to high efficiency is considered as the competitive advantage of those countries.

2.2 Coffee futures

A futures contract is a legal agreement to purchase or to deliver a specified quantity and grade of a commodity during a designated time at the predetermined price (Catania, 1989). The Coffee C futures contract is worth 37,500 pounds. The quoted price is US dollars per pound with increments of \$0.0005 per pound. Trading is performed in March, May, July, September, and December for the next 23 months. Pit trading is controlled by ICE – Intercontinental Exchange with trading floors at London, New York and Singapore. (<https://www.theice.com/products/15>). CME Globex also provides an electronic version of Coffee C futures trading.

2.3 Hedonic price model

Hedonic pricing is defined as the method of investigating the link between the price of a product and its attributes (Olanrele et al., 2014).

2.4 Efficient market hypothesis (EMH)

EMH suggests that information on the market is reflected in the price. According to Fama (1991), there are three forms of market efficiency: weak EHM proposes that past price movements are reflected in the price, medium EHM contains public information and strong one has even private information.

2.5 Price discovery

Price discovery is a specific method of pricing which assumes the balance between supply and demand, as well as their perception of the future value of some commodity (Fortenbery and Zapata, 1996). Price discovery is fairly high for futures market compared to cash market because of inherent leverage, low transaction costs and lack of short sell restrictions (Tse, 1999).

3. LITERATURE REVIEW

3.1 Introduction

Commodity futures' markets like coffee futures are volatile. Returns are high but the risks are significant. Therefore, the prices of agricultural commodity futures have been the topic of extensive research in finance (Kumar & Shollapur, 2015; Donati et al, 2016). However, most of the research is heavily mathematical and quantitative. It also focuses on commodity futures pricing in general or conducts analysis of most popular commodities like crude oil, natural gas, and gold. This paper attempts to look at the futures pricing, coffee futures in particular, from an alternative viewpoint which considers the economic features of an agricultural commodity such as weather, and past price. Coffee futures with 18 weeks to maturity, rather than longer contracts is selected because they are highly liquid and more actively traded.

This literature review will first give a short description of Arabica coffee futures and other related terms such as hedonic regression model, market efficiency, price discovery, elasticity. Insights into the criteria used to select the appropriate theoretical model as well as variables for this model are then provided. The market conditions (market efficiency and price discovery) are also discussed based on previous studies but are not tested due to the limited scope of this paper.

Dialectical approach¹ is used in this literature review to compare and contrast some highly controversial hypotheses. Moreover, methodological approach fits well with the methodology described later in the thesis. The limitation in the number of academic resources reviewed in the literature is expected due to the fact that the topic of the thesis has not been widely explored. This paper hopes to tighten the gap to some extent and to present a point of view for further study.

3.2 Background

Scholars have had different perspectives towards price forecasting accuracy. Some studies refute its positive impact due to the high level of bias, the lack of useful information and the inefficiency in predicting the price (Mohan & Love, 2004; Krichene, 2008). On the other hand, several researchers support the potential usefulness of price

¹ Dialectical approach is a research method aiming to analyze and compare contrasting perspectives

forecasting, especially for commodity futures market (Ajao, 2013; Rauch et al. 2013; Lee, 2015).

There have been concerns regarding the oversimplified assumptions of predicting future price using past time-series data (Bernake, 2008; Levine et al., 2010). It is claimed that price forecasting depending solely on past data was of high subjectivity because the future price might also react strongly to other external factors, such as supply and demand on the market, trader's personal decision-making, and technology. However, price forecasting is considered to assist in reducing uncertainty in the market. Speculators also use it to achieve profit-seeking goals by risk-taking while hedgers participate in commodity futures market to protect physical goods from price fluctuations. Both of them perceive price prediction to have close links to profitability (Schwager, 1995; Aguiar & Borestein, 2012; Pinheiro & Valter de Senna, 2016).

At the annual meeting in 2008, Bernanke, the Chairperson of the Federal Reserve Bank of Boston, cast doubt on the level of accuracy of commodity futures pricing, especially for the government in terms of social change and welfare gain or loss. Many types of commodities like energy and food account greatly for export revenues or are closely related to economic development. As a result, fresh methods of forecasting commodity futures price are encouraged to be presented.

The thesis aims to provide a way to fill the aforementioned gap using an integration between traditional financial prediction and economic elements like weather, supply and demand, and past price. It combines both past time-series data and present data to minimize the limitations addressed in previous studies.

3.3 Previous studies of coffee futures pricing

Most research papers focus on a wide range of commodity futures or a particular type of commodity futures; however, coffee futures has not been widely studied. A variety of models have been extensively introduced and reviewed for pricing commodity futures. Two of the most famous are the models by Black (1976) and Schwartz (1997). Both of them are one-factor models which propose that "futures price as a function of spot price and the time to maturity" (Rauch et al., 2013: 144). However, there still are

some limitations on those models because the spot price is not easily modeled by the Black equation and “modeling potential is limited to elementary characteristics” (ibid).

The two-factor model by Trolle & Schwartz (2009) which added convenience yield and different multi-factor models by Cotazar et al. (2008), Paschke and Prokopczuk (2010), Dempster & Tang (2011) are proved to give better results for a single commodity market than for a commodity portfolio. Nevertheless, in regards to the concerns of Bernake (2008), this research attempts to bring a new perspective to this topic by focusing on only coffee futures. It is generally believed that futures price reflects the price of the cheapest deliverable grade. In contrast, the empirical research of Kuhn (1988) revealed that coffee futures price reflects cash price of the most widely traded and deliverable grade of coffee.

While previous researches mainly used data from the same market to estimate commodity future contracts, Chen et al. (2008) and Ajao (2013) presented formulae to predict commodity futures price based on Forex futures in the former and on exchange rates in the latter. These studies illustrate that such method might be better than traditional statistical methods like auto-regression. Forecasts are also more efficient thanks to the availability of information. Exchange rates could be obtained and updated frequently while economic data used in auto-regression is not often revised. However, the accuracy of a model depends substantially on the type of commodity traded, and prediction with exchange rate might overgeneralize the distinct characteristics of each market. For instance, Chen et al. (2008) selected data from Australia, Canada, Chile, New Zealand, and South Africa since primary products generate a large amount of export revenues for those countries. Nevertheless, Brazil and Colombia are two main players of Arabica coffee exports, and their exchange rates (Brazilian Real and Colombian peso) are not related to above data. Therefore, coffee futures cannot be applied to this case.

3.4 Market efficiency

The market efficiency hypothesis is significantly controversial. If all available information is incorporated into market prices, traders cannot beat the market. Therefore, technical analysis, which is the technique of using past price and linked statistics to predict price movements, violates the widely believed efficient market

hypothesis (Ghobadi & Abdolbaghi, 2014). On the other hand, fundamental analysis, which might not perform correctly if the market has the strong or semi-strong form of efficiency, is still more supported and accepted if the market is weakly efficient, or inefficient.

Market efficiency, especially for commodity futures, varies regarding the location and type of commodity. Di Matteo et al. (2005) cited in Kristoufek & Vosvrda (2013) have suggested that developed markets like the USA and Japan are more efficient than developing ones. Moreover, researchers have reported different results regarding the level of efficiency of the coffee futures market. Empirical analysis of Mohan & Love (2014) supported the inefficiency of the coffee futures market. If all information that informed traders collect is reflected in the price, other traders can simply observe the price of coffee futures and obtain the information free. Both Kofi (1973) and Kebede (1992) imply that the short-term market (1-33 weeks of maturity) had a weak form of efficiency for coffee futures. Kristoufek and Vosvrda (2013) ranked 25 commodity futures markets in terms of their efficiency, and coffee was in the top five. It is worth noticing that this study's scope is only among commodity futures. The lack of comparison with other financial markets should be taken into consideration. As a consequence, the strong form of efficiency of coffee futures is not ensured.

Overall, the thesis assumes that coffee futures market, especially nearby maturity contracts, has the weak form of efficiency. It is because there are cases that traders can beat the market. In this case, fundamental analysis which will be conducted later could provide more reasonable results.

3.5 Price discovery

Gross et al. (2016) cite Working (1962) and Black (1976) when making reference to the facilitation of price discovery in futures trading. Theoretical arguments suggest that because futures markets have higher liquidity, more transparency, and lower transaction costs, they are more sensitive to latest information than spot markets. Keat & Young (2006) also claim that traders in the market of commodity futures tend to find the most accurate price estimator to fulfill price discovery. Furthermore, liquidity plays an important role in price discovery (Chatrath & Christie-David, 2004). Therefore, 18-week Coffee C Futures might be a suitable object of price discovery.

3.6 Methodology: hedonic regression model

Knowledge of the factors that have an effect on futures pricing can also be considered as fundamental analysis. This is one of the most popular techniques that traders use to seek returns in the futures market. The origin of the technique can be traced to “forecasting theory” which proposes that returns from futures trading in markets like coffee rely highly on the ability of the traders to correctly predict price (Lee J., 2015).

In the previous studies, hedonic regression model has been often applied to real estate industry or only spot price of a commodity (Djunaidi, 1993; Olanrele, 2014). Regarding the choice of variables, Schwager (1995) defines fundamental analysis as the use of economic data (production and consumption) to predict price. Those previous studies spur this selection of hedonic regression model.

3.6.1 *Weather*

Coffee futures experiences high volatility partly because weather, a crucial factor, is unpredictable. Otero & Milas (2001) included weather conditions of countries producing coffee as variables in their coffee price modeling. They might cause short-term effects on the price of coffee (exogenous shock). In practice, weather has prime importance to commodities. Adverse weather could ruin coffee production, while favorable weather can result in good harvest. It simultaneously affects quality of coffee, as well as imports and exports.

3.6.2 Exchange rate

A primary product like coffee tends to have supply and demand inelasticity (Otero & Milas, 2001), which means that a move in one curve might result in a greater change in price. Therefore, information on supply and demand of coffee could have a considerable effect on the price of coffee futures. Supply and demand are subjected to the change in the volatility of exchange rate, especially commodities like coffee futures.

Trade flows are influenced by the unexpected movement of exchange rates. It is still a debate that whether the variability of the exchange rate has a positive or negative

impact on international trade. In particular, the fact that exchange rates are highly fluctuating leads to the rise in global trade. However, others believe that the volume and value of international transactions are negatively affected due to higher risks (Susanti, 2001; Ajao, 2013).

3.6.3 Spot price and futures price

A large number of studies on futures market and finance efficiency market have different perspectives towards the association between spot price and futures price. Turnovsky (1983) and Choe (1990) both demonstrated that futures price has a close link with commodity prices in the futures. However, Wang and Ke (2002) refuted the hypothesis that futures price could be an unbiased predictor of the product price in the future. Their research resulted in the underperformance of futures price to forecast subsequent futures spot price.

Coffee futures price is likely to change accordingly to spot price (Mohan & Love, 2004). Bulk cocoa prices were assumed to be closely tied to the futures price with the coefficient of 0.68 (Susanti, 2001). Coffee spot price, similarly, could be identified as a factor affecting price of coffee futures.

3.6.4 Autocorrelation of price

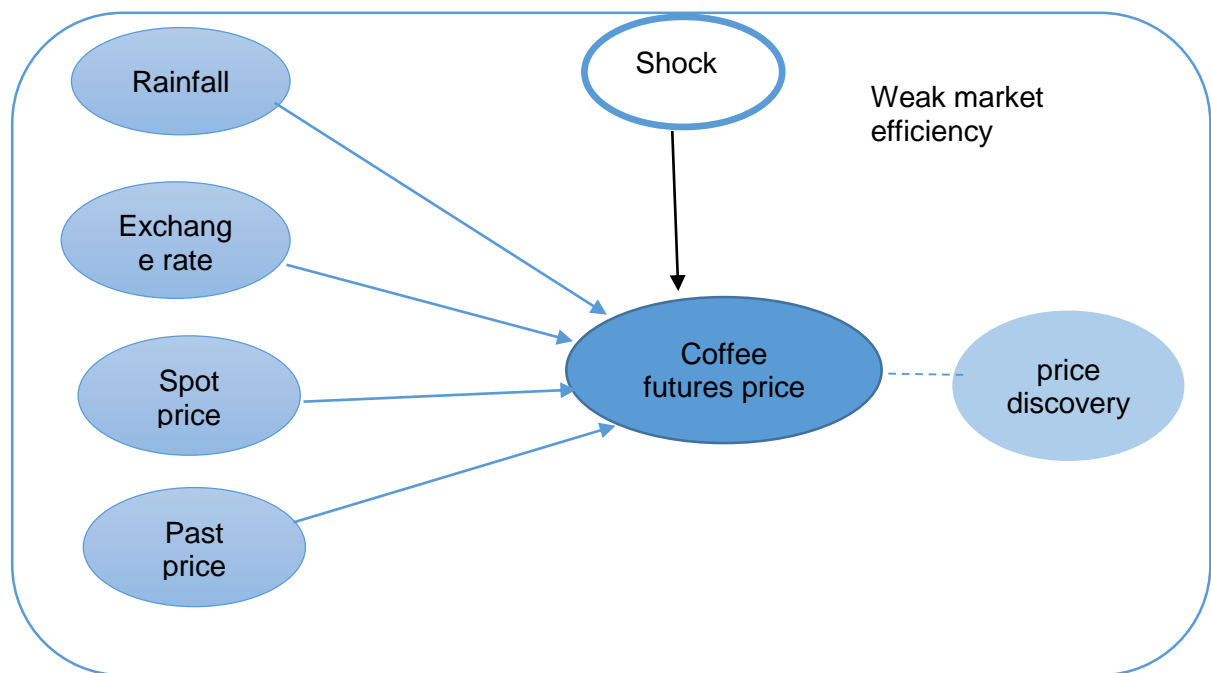
Autocorrelation of commodity prices is characterized by small kurtosis², positive autocorrelation, and positive skewness (Tomek, 2000; Mitra & Boussard, 2012).

3.7 Conclusion

This paper, by empirical research, attempts to provide an evidence to the popular debate: the possibility of the efficiency of fundamental analysis in predicting commodity futures price, particularly coffee futures. Price forecasting is not only important to traders who seek for high returns but also to the economy in general. Other players in the market such as producers/manufacturers, consumers, governments might also use futures price as one of the guidance for allocating the resources or making purchasing decisions.

3.8 Conceptual framework

² Kurtosis: the statistical indicator to measure the sharpness of the peak of the frequency-distribution curve



4. METHODOLOGY

4.1 Theoretical model

The hedonic price model is based on a combination approach called “hedonic function”. It considers attributes as inputs to test the model, which might provide the output as the market price of a product (Coulson & Robin, 2001), $P = f(X_1, \dots, X_n)$. Lucas (1974) cited in Djunaidi (1993) proposed that hedonic pricing is a regression of the observed prices of a commodity and the characteristics related to them.

Espinosa and Goodwin (1991) have used hedonic pricing for Kansas wheat. Djunaidi (1993) studied Indonesian cocoa price with that model as well. Usually, hedonic regression pricing supports studies of property markets, for example, Peterson & Flanagan (2009) and Olanrele et al (2014).

The theoretical model for this thesis has been reviewed by Djunaidi (1993) and Oranrele et al (2014). Define the following:

$$P = f(X) \text{ or } P = f(X_1, \dots, X_n)$$

where X ranges from X_1, \dots, X_n

Considering $f(X)$ as a linear function:

$$P = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n$$

In the linear relation, using calculus differential equation,

$$\partial P / \partial X_1 = a_1$$

A change in P due to change in X_1 is constant and equal to the coefficient β_1 .

Adding the possibility of least square errors, the hedonic regression equation could be written as:

$$P = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n + \varepsilon$$

Where X_1, \dots, X_n are the characteristics of the product and P is the price of the product. Statistical analysis is necessary for the function to calculate the values of β which give the influence of each X on the price.

4.2 Empirical model

The thesis uses statistical tests (mainly hedonic regression model with time series data) to analyze the effect of several factors to coffee futures price because there is a belief that futures price might react quickly to information on factors associated with the physical price of a commodity.

Djunaidi (1993) used a unit root test, Johansen's cointegration test and Granger's causality test for forecasting Indonesian daily cocoa prices. Kumar & Shollapur (2015) conducted unit root test, Johansen's cointegration test and vector error correction model to study price discovery in agricultural commodity futures market in India. However, this quantitative analysis is subject to limitations due to the narrow scope of a bachelor's thesis, as well as budget and analytical tool constraint. Variables are chosen partly because of their availability.

The unit root test was first used for times series variables, then other tests for multiple linear regression assumptions were conducted: normality, linearity, independence and

homoscedasticity. A correlation matrix was formed between variables. Multicollinearity (whether independent variables are independent of each other or not) was checked before multiple linear regression is applied. Period of time for the 228 samples is from January 1994 to December 2012 (19 years). Monthly data is exploited because the model contains many short-term effect factors, which are not well presented under daily, quarterly or annual term. Excel and *x/stat* Add-in is used as the analytical tool.

4.3 Variable selection

4.3.1 Independent variables

The selection of independent variables is derived from one of the two most popular methods of evaluating a security: fundamental analysis (apart from technical analysis). Macroeconomic indicators and physical attributes related to coffee commodity are under consideration.

First precipitation in Brazil and Colombia, the two largest Arabica producers are collected from Climate Change Knowledge portal of World Bank. Because Coffee Futures is traded in the base currency of U.S dollar internationally, the exchange rate of Brazilian Real and Colombian Peso over U.S dollar are included in the empirical model. Yin (2001)'s findings lead to the claim that commodity futures form a relationship with commodity prices. Therefore, the spot price of Arabica coffee is captured for the purpose of examining whether commodity futures market has a close link with spot market, so that information having an impact on spot market also affects futures market simultaneously. The data of those three variables are from International Financial Statistics dataset of International Monetary Fund. Based on the assumption of the auto-correlation between futures prices, Arabica futures price of the previous month (P_{t-1}) is hypothesized in the regression. Historical Coffee C Futures data is retrieved from Intercontinental Exchange (ICE Futures U.S).

4.3.2 Dummy variable

Coffee futures experiences some exogenous shocks that affect the price series during the examined sample time. As a result, the dummy variable S representing shocks is added to the regression model. According to ICO, from the second half of 1994 to first half of 1995, adverse weather conditions in Brazil caused a significant rise in the price of coffee export. That could correspondingly increase Arabica futures price. The report

named 'Lessons from the world coffee crisis: A serious problem for sustainable development' from ICO (2004) expressed the severe and negative results of Coffee Crisis from 1999 to 2004. Arabica futures during that period experienced a substantial drop in price. Financial crisis in 2007-2008 and the later economic depression are also been into account. Currency movements were considered as "the driving force in the price behavior of futures market for many commodities, and coffee is no exception" (ICO, 2009: 3). International coffee outlook at that time underwent instability. Dummy variable S takes the value of one during those above mentioned periods of time, while takes the value of zero in other cases.

4.3.3 Dependent variable

Historical Coffee C Futures data is retrieved from Intercontinental Exchange (ICE Futures U.S). The original form of data is the daily price from January 3, 1994 to December 31, 2012. However, they were converted into average monthly price series for the purpose of the thesis.

4.4 Unit root test

Stationary in time series data is based on the assumptions that the series have unchanged statistical mean, variance or correlation. Stationary helps to predict the future considering the stability of the series. The time series data usually faces non-stationarity, which can mislead regression model to spurious results. Non-stationarity might be due to deterministic trends, cycles, random walks or the mix of three. To diagnose whether time series data used in this thesis is stationary or not, Augmented Dickey-Fuller Test (ADF Test) for Unit Root at 5% significance level is performed. It is to ensure that there is no unit root property. In addition, first differencing is necessary for the data that might have unit root. First differencing is subtracting Y_{t-1} from Y_t .

4.5 Suggested empirical coffee futures model

$$P_t = \beta_0 + \beta_1 PRE_{Br} + \beta_2 PRE_{Co} + \beta_3 USD/BRL + \beta_4 USD/COP + \beta_5 SPOT + \beta_6 P_{t-1} + \beta_7 S + \varepsilon_t$$

Defining the following:

P_t : Price of Arabica coffee futures, monthly average

β_0 : Intercept

PRE_{Br}: Precipitation of Brazil

PRE_{Co}: Precipitation of Colombia

USD/BRL: Exchange rate of U.S dollar and Brazilian Real, monthly average

USD/COP: Exchange rate of U.S dollar and Colombian Peso, monthly average

SPOT: Spot price of Arabica coffee

P_{t-1}: Coffee futures price of the previous month in the time series

S: Dummy variable for exogenous shocks

ε_t : error term

Several tests should be conducted in order to ensure the validity of the multiple regression model. Those are based on four multiple regression assumptions (Wooldridge, 2012; Olanrele et al, 2014):

1. Normality of distribution: Non-normality of distribution of the residuals might mislead confidence intervals for model forecasting capability. It could cause error distribution to be skewed by the appearance of several outliers.
2. Linearity: The dependent variable and independent variables should have linear relationship
 - 2.1 The expected value of dependent variable is a straight-line function of each independent variable, holding the others fixed.
 - 2.2 The slope of that line does not depend on the values of the other variables.
 - 2.3 The effects of different independent variables on the expected value of the dependent variable are additive.
3. Correlation among variables: Each independent variable should not be highly correlated with each other (no multicollinearity).
4. Homoscedasticity: The variance of error term should be constant across time for time series data and across independent variable.

5. FINDINGS

5.1 Unit root test

Among the variables, only PRE_{Br} and PRE_C are stationary.

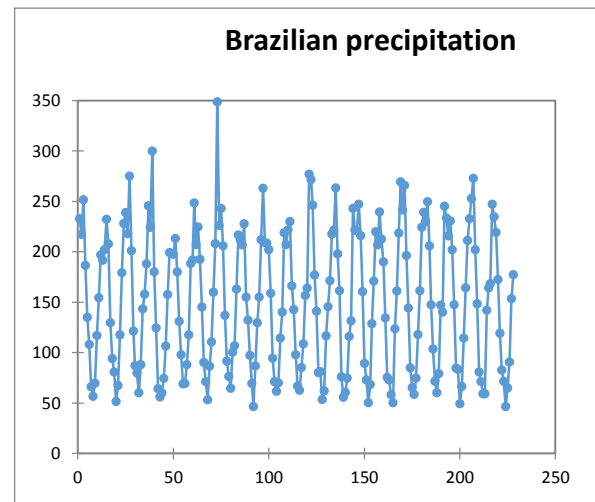
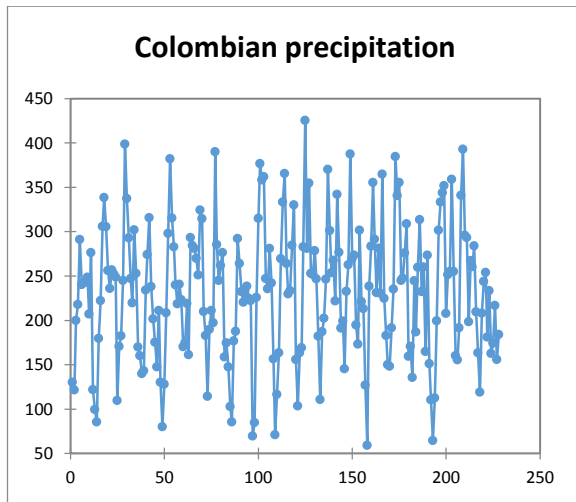
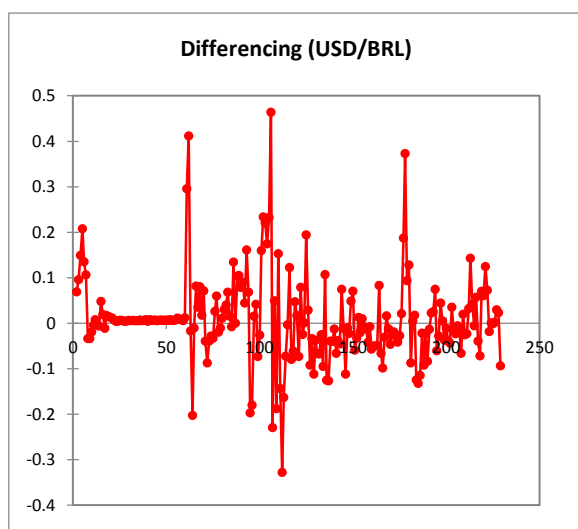
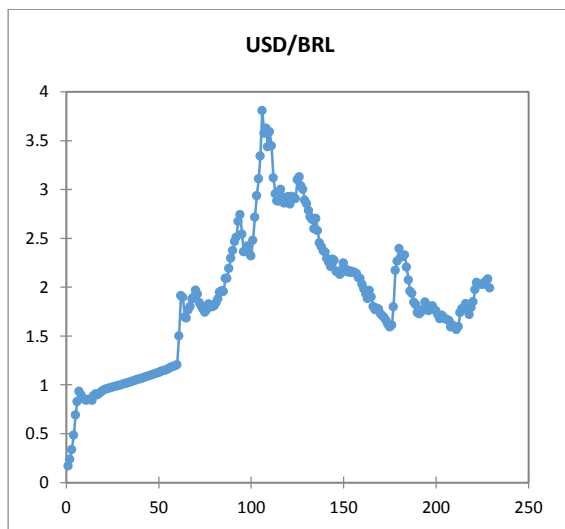
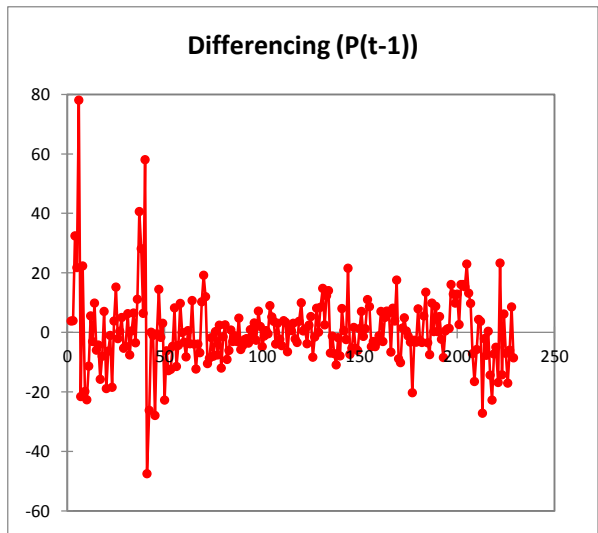
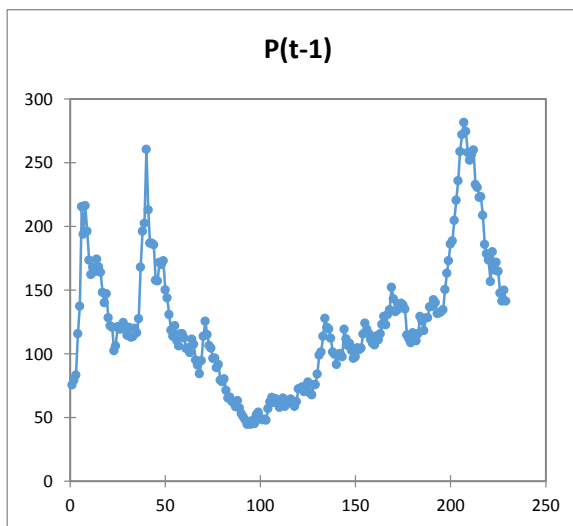
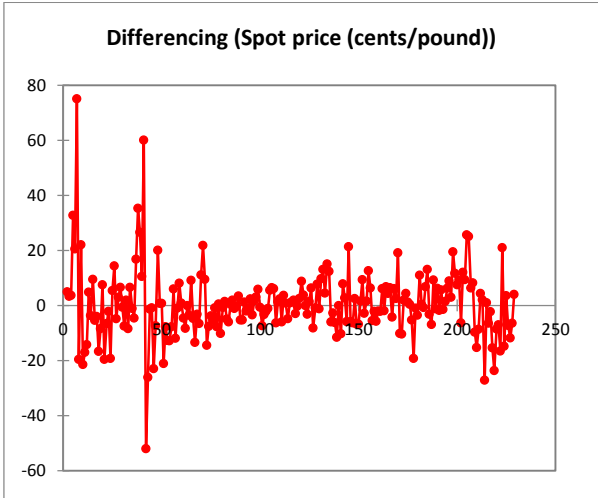
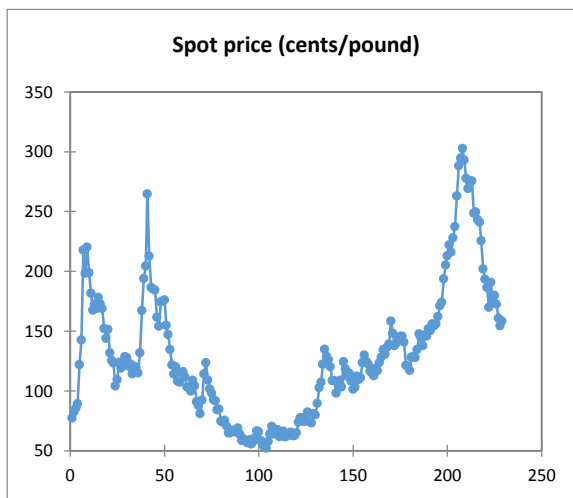
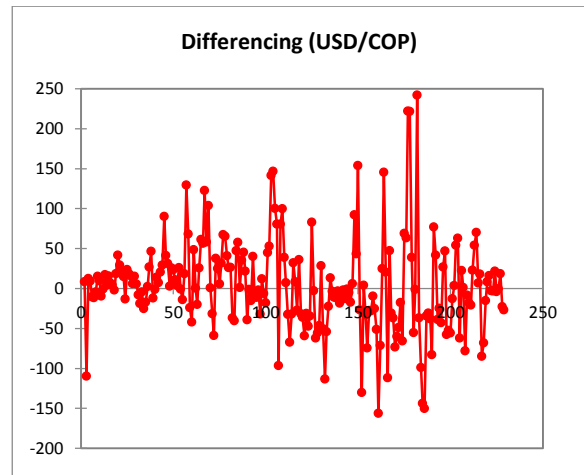
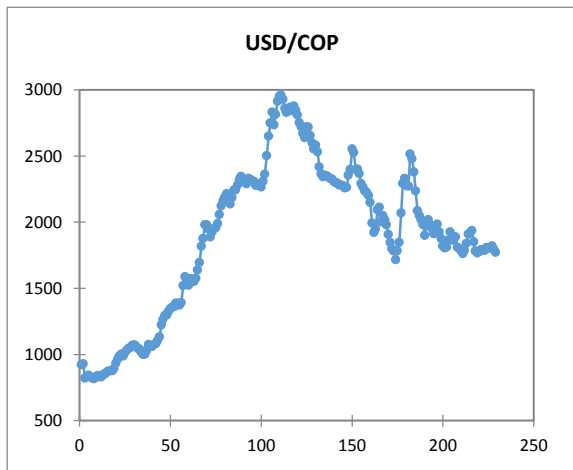


Figure 1: Plots of the levels of PRE_{Br} and PRE_{Co}

At 95% confidence, all other time series variables exhibit non-stationary behavior. Therefore, first differencing is necessary for further modeling.





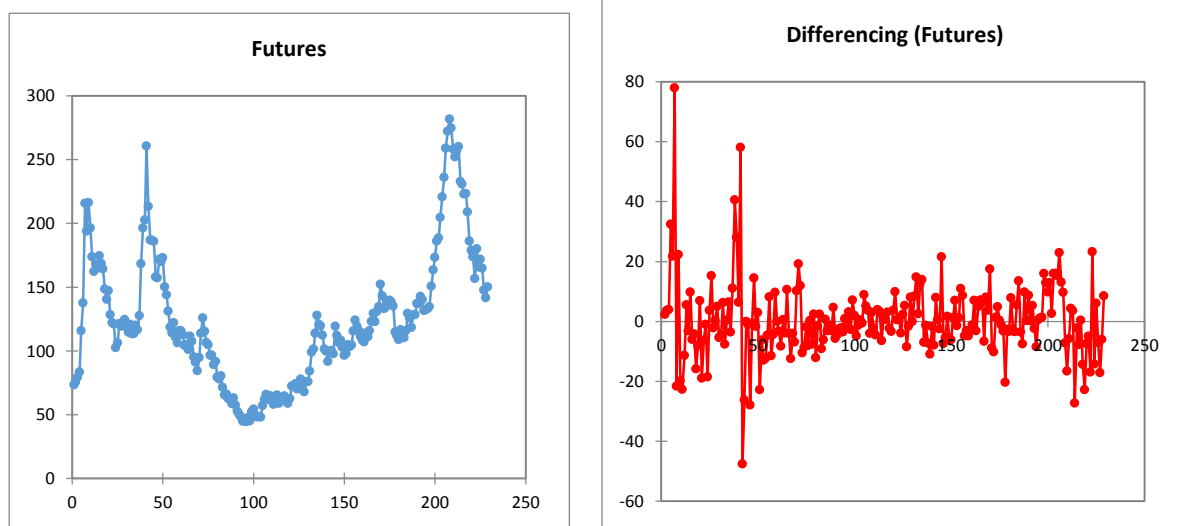


Figure 2: Plots of the levels and first differences of the other independent variables

5.2 Tests for multiple regression assumptions

5.2.1 Normal distribution test

The residual is calculated as the error (difference) between observed values (observed coffee futures prices) and predicted values.

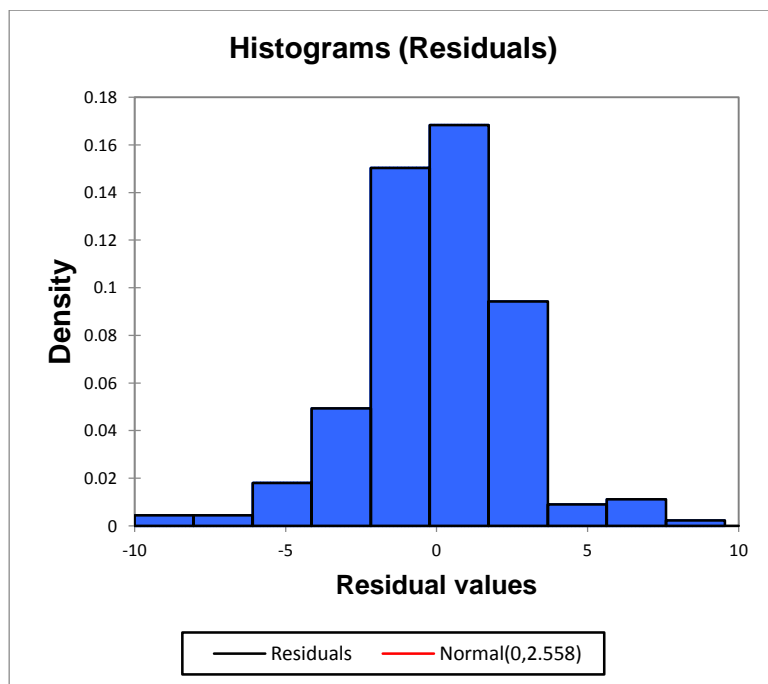


Figure 3: Histogram of residuals of the regression

The histogram of residual values of the regression with a fitted normal curve shows that they are roughly normally distributed.

5.2.2 Linearity test

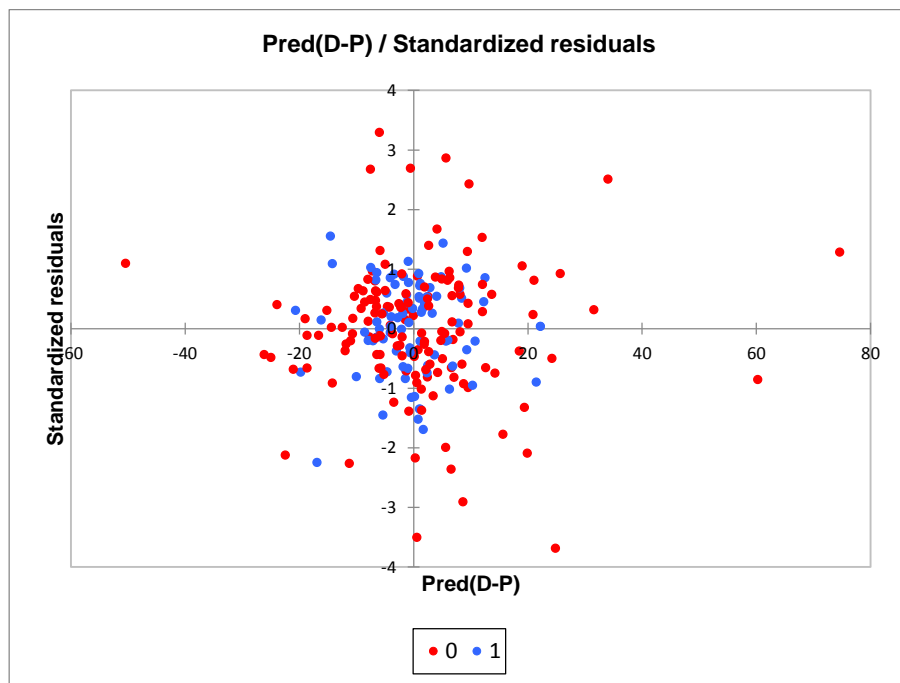


Figure 4: Predicted DP_t versus residuals plot

The residuals scatter roughly symmetrically around the predicted coffee futures price line (DP_t), which could imply the linearity of the regression model.

5.2.3 Correlation tests

	PRE_{Br}	PRE_{Co}	DUSD/BRL	DUSD/COD	DSPOT	DP_{t-1}
PRE_{Br}	1	-0.507	-0.086	-0.149	0.125	0.069
PRE_{Co}	-0.507	1	0.050	0.069	-0.102	0.043
DUSD/BRL	-0.086	0.050	1	0.449	-0.038	0.038
DUSD/COD	-0.149	0.069	0.449	1	-0.164	-0.057
DSPOT	0.125	-0.102	-0.038	-0.164	1	0.132
DP_{t-1}	0.069	0.043	0.038	-0.057	0.132	1

Table 1: Pearson's correlation matrix between independent variables

Most of the correlation values are under 0.09, which implies that those independent variables are not correlated to each other. However, exchange rate USD/BRL and USD/COD have the moderate correlation of 0.449 because both currencies are from the same continent, and both countries Brazil and Colombia might face similar change in economy and finance.

The close link between spot price and futures price that has been discussed in previous parts of the thesis could partially lead to the correlation value 0.132 between SPOT and P_{t-1}

The correlation between the precipitation in Brazil (PRE_{Br}) and the precipitation in Colombia (PRE_{Co}) is not as high as expected (-0.507).

	PRE_{Br}	PRE_{Co}	DUSD/BR L	DUSD/CO D	DSPO T	DP_{t-1}
Tolerance	0.720	0.733	0.793	0.765	0.944	0.965
VIF	1.389	1.365	1.261	1.307	1.060	1.036

Table 1 Multicollinearity statistics

The tolerance values of all independent variables are greater than 0.2. The Variance Inflation Factor values are all smaller than 10. Therefore, no multicollinearity exhibits in the independent variables.

5.2.4 Homoscedasticity test

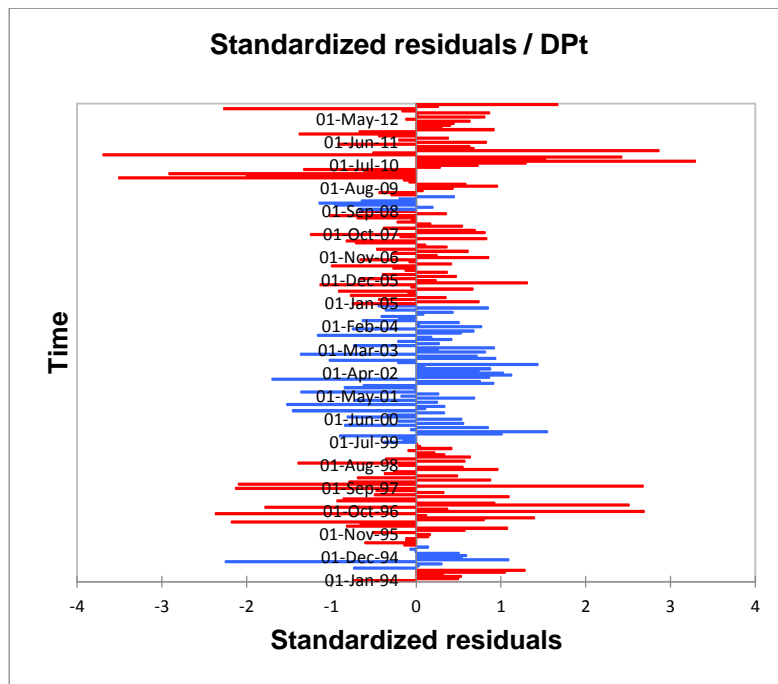


Figure 5 Standardized residuals versus time plot

Figure 4 Predicted DP(t) versus residuals and Figure 5 Standardized residuals versus time plot suggests that variance of the residuals exhibit consistency and the residual of the dependent variable is not affected by the independent variables.

5.3 Descriptive statistics

Statistic	PRE _{Br}	PRE _{Co}	DUSD/BRL	DUSD/COD	DSPOT	DP _{t-1}	D-P
Number of observations	228	228	228	228	228	228	228
Number of missing	0	0	0	0	0	0	0
Minimum	46.377	59.204	-0.328	-156.406	-51.944	-47.562	47.562
Maximum	348.746	425.267	0.464	241.962	75.107	77.985	77.985
Median	147.896	235.453	0.005	0.524	-0.511	-1.033	-0.790
Mean	150.902	232.114	0.008	3.724	0.356	0.288	0.336
Standard deviation	68.372	75.196	0.092	56.744	12.305	12.355	12.341
Skewness	0.170	0.039	1.136	0.737	1.283	1.408	1.402
Kurtosis	-1.099	-0.330	5.551	3.058	8.932	9.266	9.296
Standard error (Skewness)	0.161	0.161	0.161	0.161	0.161	0.161	0.161
Standard error (Kurtosis)	0.321	0.321	0.321	0.321	0.321	0.321	0.321

Table 2: Statistics for normal distribution test

PRE_{Br} and PRE_{Co} both have skewness and kurtosis fitting in the acceptable range of roughly normal distribution between -1.96 and 1.96. Other variables have acceptable skewness but very high kurtosis.

5.4 Hedonic regression model

The data were entered into Excel xlstat to run for multiple linear regression following the Hedonic pricing model. The regression provides the intercept β_0 and the coefficients to form the following equation:

$$P_t = -0.787 - 0.003PRE_{Br} + 0.004PRE_{Co} - 1.607DUSD/BRL + 0.002DUSD/COD + 0.987 DSPOT - 0.019DP_{t-1} + 0.411S$$

Given the adjusted coefficient of determination $R^2 = 95.6$, nearly 96% of the variability of Coffee futures price is explained by the seven independent variables. Other independent variables that are not put into the hedonic regression function might also explain the rest of coffee futures price.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	7	33088.162	4726.880	700.179	< 0.0001
Error	220	1485.211	6.751		
Total	227	34573.373			

Table 3 Analysis of variance

The error of the model is 1485.2, and the mean square error of 6.8. The Fisher's F test is applied to determine whether the used regression model fits the population from the dataset. The probability corresponding to the F value is lower than 0.0001 at the significance level of 5%. The result demonstrates that the null hypothesis (no effect of the two explanatory variables) is assumed to be incorrect at a lower than 0.01% risk. Therefore, it might lead to the confirmation that the seven independent variables do provide a considerable amount of information.

Source	Value	Standard error	T	Pr > t	Lower bound (95%)	Upper bound (95%)
Intercept	-0.787	0.970	-0.811	0.418	-2.698	1.125
PRE _{Br}	-0.003	0.003	-0.883	0.378	-0.008	0.003
PRE _{Co}	0.004	0.003	1.479	0.141	-0.001	0.009
DUSD/BRL	-1.607	2.099	-0.765	0.445	-5.743	2.530
DUSD/COD	0.002	0.003	0.524	0.601	-0.005	0.009
DSPOT	0.987	0.014	68.431	< 0.0001	0.959	1.016
DP _{t-1}	-0.019	0.014	-1.325	0.186	-0.047	0.009
Schock	0.411	0.358	1.149	0.252	-0.294	1.116

Table 4 Unstandardized regression coefficient

At the confidence level 95%, only Spot price has significant contribution in the model. Other variables such as precipitation in Brazil and in Colombia, exchange rate of USD over Brazilian Real and Colombian Peso, past coffee futures price, and even shock are not significant (0.378, 0.141, 0.455, 0.601, 0.186 and 0.252 respectively).

6. DISCUSSION AND ANALYSIS

Among the coefficients, spot price has the highest coefficient (0.987), which might demonstrate that Arabica futures price from January 1994 to December 2012 has a close relationship with spot price. However, spot price, which is not usually considered as a fundamental factor in fundamental analysis, has raised predicted value (R^2) of this hedonic regression model significantly. High coefficient further suggests that coffee

futures move nearly similarly to coffee as commodity price. Hedgers, usually farmers and producers, who aim at reducing commodity price risks, could take advantage of the association between spot prices and futures prices. The result conflicts with some of the previous researches which have undermined the possibility of price discovery. Regulatory bodies, such as policymakers and international organizations, especially from countries heavily depending on exports could also partly observe the movement direction of futures prices for subsequent spot prices.

Precipitation in two main Arabica coffee producers Brazil and Colombia generates opposite coefficients. This result goes against the intuitive assumption of similarity between two variables. Rainfall in Brazil has negative coefficient with Arabica futures price, which implies that the price will increase when precipitation level decreases. The opposite case is true for Colombian precipitation. Moreover, the coefficients are not of high significance. It might be due to the low frequency of data analyzed (monthly), when weekly or even rainfall could have a short-term effect on the growing process of coffee. Usually, daily futures price could capture that type of information quite simultaneously, while monthly data could not. Another argument is the differences in geographical location of the two countries concerned. They are on the different longitude and surrounded by Pacific Ocean (Colombia) and Atlantic Ocean (Brazil). Rainfall data obtained are for the whole nations but not specifically focus on regions producing coffee.

The exchange rate USD/BRL and USD/COL also have contrasting coefficients with coffee futures price. Exchange rate USD/COL could be said to move in the same direction with coffee futures price, while USD/BRL does not. USD/BRL increased dramatically about ten times during the observed period of time, but USD/COL only rose two times. The devaluation of local currencies, which leads to the rise in the exchange rate of USD over those currencies has matched with growing coffee production in Brazil and Colombia. That USD/BRL was highly volatile could be because Brazil is dependent on foreign investment to finance growth. Brazilian Real experienced great devaluation in the observed period of time. On the other hand, the more stable condition of Colombian peso might result from the improvement in its

economy (ICO, 2014). USD showed stability over time. The small coefficients of two exchange rates might change in later time because the impact of Brazilian Real and Colombian Peso on the international coffee market tends to increase as the market share of two countries grow consistently (Cashin P. et al., 2003).

With the negative coefficient (-0.019), the coffee futures seems not to have autocorrelation with its nearby past price. It could be interpreted as the result does not support several studies' conclusion that commodity futures prices are auto-correlated. However, the notion that only the futures price of the past month is included in the regression model should be taken into consideration. The coffee price might have a seasonal or periodical pattern that could not be detected in the scope of this paper. Past price at different time frame might provide better results.

The coefficient of shock is 0.411. One possible explanation is that when there is a shock (usually a crisis or severe weather conditions in countries producing Arabica), the coffee price, and coffee futures price move up due to the shortage of supply.

7. CONCLUSION

7.1 Main findings

Unit root test is conducted for 228 samples of Coffee "C" Futures and the independent variables: precipitation in Brazil and in Colombia, exchange rates USD/BRL and USD/COD, spot coffee price and coffee futures price of the previous month from January 1994 to December 2012. The Augmented Dicker-Fuller test provides the results that only two precipitation variables are stationary while the others could have unit root at 95% confidence interval. The first differencing method is used to remove unit root in those time series data.

Several tests for ensuring multiple regression assumptions are performed. Normal distribution histogram for residuals confirms the normality assumption. Pearson's correlation matrix and multicollinearity test could rationalize the no-collinearity assumption. The scatterplot of predicted coffee futures price (dependent variable) versus residuals implies the linearity of the model. The plot of residuals over time refutes the heteroskedasticity of the time series data. The input variables could be said to be valid for conducting hedonic regression model.

The hedonic regression model reveals that 96% of coffee futures price could be explained by precipitation in Brazil and in Colombia, exchange rates USD/BRL and USD/COD, spot coffee price, coffee futures price of the previous month and dummy variable shock. Only spot price is of great significance at 95% confidence, while the others are not. Spot price also has the highest coefficient, while coefficients of other independent variables are relatively low.

The simplicity of the hedonic pricing model could be an advantage because it directly provides a useful forecast. Traders could obtain information on the above fundamental factors as a guide to predict the movement of coffee futures. Farmers and producers can plan their operations accordingly to the market with more confidence.

7.2 Limitation of the research

The short period of time series data (January 1994 to December 2012) and the inability to acquire most recent data might affect negatively on the accuracy and long-term usability of the model. The scope of the thesis is limited with only 228 samples and low frequency. Most independent variables are monthly data, which requires other variables to be converted to the same frequency.

The hedonic regression model is based on past movement of both coffee futures and fundamental factors, as well as historical shocks. It is not ensured that the same pattern

would be appropriately applied for future movement. Forward-looking shocks, especially dramatic ones might be difficult to be diagnosed.

The independent variables are mainly in supply side (precipitation of coffee production countries and exchange rate of those countries over U.S dollar). Moreover, Brazil and Colombia do not account for 100 per cent Arabica production. Other supply countries such as African countries and Indonesia and demand countries such as the United States and European countries have not been explored in the research.

7.3 Suggestion for further research

The hedonic regression pricing for coffee futures could be later modified for other commodity futures (cocoa, sugar, grains). They react quickly to weather, supply and demand, and exchange rate information. The commodities are also export and import products for many countries. Daily data are suggested for regression model because they capture more effectively the influence of fundamental factors. Additionally, speculators also prefer short-term information to earn profits in short period of time.

More advanced analysis tests and models such as co-integration test, causality test, and error correction model are needed to improve the performance of the hedonic regression model, as well as the validity of modeling results. The exchange rate of USD/EURO could be taken into account in further research due to the high level of contribution of European countries to international coffee demand.

Reference list

- Adämmer P. Bohl M. & Gross C. (2016) 'Price Discovery in Thinly Traded Futures Markets: How Thin is Too Thin?' *Journal Of Futures Markets*, 36(9): 851-869.
- Ajao Y. (2013) *Performance prediction of commodity prices using foreign exchange futures*. Doctoral dissertation. Colorado, USA: Walden University [published]
- Bernanke, B. (2008) *Financial regulation and financial stability*. Speech delivered at the Federal Deposit Insurance Corporation's Forum on mortgage lending for low- and moderate-income households. Arlington, VA.
- Black F. (1976) 'The pricing of commodity contracts.' *Journal of Financial Economics*, 3: 167–179.
- Cashin P., Céspedes L., & Sahay R. (2003) 'Commodity currencies and the real exchange rate'. *Journal of Development Economics*, 75: 239-268.
- Catania, P. J. (1998): *Commodity Trading Manual*. Chicago: Chicago Board of Trade.
- Chatrath A, & Christie-David R. (2004) 'Futures Expiration, Contract Switching, and Price Discovery.' *Journal Of Derivatives*, 12(1): 58-72.
- Chen Y.-C., Rogoff K., & Rossi B. (2008) 'Can exchange rates forecast commodity prices?' In: *NBER Working Paper No. 13901*. Cambridge, MA: National Bureau of Economic Research.
- CME Coffee C futures quote*. Available from <https://www.theice.com/products/15> [Accessed February 2 2017]
- Cortazar G., Milla C., & Severino F. (2008) 'A multicommodity model of futures prices: Using futures prices of one commodity to estimate the stochastic process of another.' *Journal of Futures Markets*, 28(6): 537–560.
- Coulson, N. E., & Robin, M. L. (2001) 'The Internal and External Impact of Historical Designation on Property Values.' *Journal of Real Estate Finance and Economics*, 23: 113–124.
- Dempster M., & Tang K. (2011) 'Estimating exponential affine models with correlated measurement errors: Applications to fixed income and commodities.' *Journal of Banking and Finance*, 35(3): 639–652.
- Di Matteo T., Aste T., Dacorogna M. (2005) 'Long-term memories of developed and emerging markets: using the scaling analysis to characterize their stage of development.' *Journal of Banking & Finance*, 29: 827–851.

Djunaidi H. (1993) *Price analysis of the cocoa market: Causality testing, price forecasting, and hedonic pricing*. Ph.D. Thesis. Oklahoma, USA: Oklahoma State University [published]

Donati, M., Zuppiroli, M., Riani, M. and Verga, G. (2016) 'The impact of investors in agricultural commodity derivative markets' *Outlook on Agriculture*, 45(1): 25-31. Retrieved from Sage Journals Database [Accessed 14 Dec. 2016].

Espinosa J. & Goodwin B (1991) 'Hedonic Price Estimation for Kansas Wheat Characteristic'. *Western Journal of Agricultural Economics*, 16:72-85.

Fama, E. (1991) 'Efficient Capital Markets: II.' *Journal Of Finance*, 46(5): 1575-1617.

Futures Markets'. *The Journal of Futures Markets*, 19(8): 911-930.

Ghobadi M. and Abdolbaghi A. (2014) 'PROFITABILITY OF TECHNICAL ANALYSIS INDICATORS TO EARN ABNORMAL RETURNS IN INTERNATIONAL EXCHANGE MARKETS.' *Theoretical & Applied Science*, 19(11): 20-26.

International Coffee Organization (2004) 'Lessons from the world coffee crisis: A serious problem for sustainable development.' In: *UNCTAD XI Conference*; São Paulo, Brazil: 13-18 June. London: ICO. Pp 2-3.

International Coffee Organization (2009) 'The world economic crisis and the coffee sector.' London: ICO. Pp.3-5.

International Coffee Organization (2014) 'Coffee and currencies: How the buck impacts the bean'. London: BNP Paribas.

International Trade Centre (2004) *Bitter or better future for coffee producers?* Retrieved from Business Source Complete, EBSCOhost [Accessed 5 January 2017].

Keat P., & Young P. (2006) *Managerial Economics-Economic Tools for Today's Decision Makers*. New Jersey: Prentice Hall.

Kebede Y. (1993) 'Causality and Efficiency in the Coffee Futures Market'. *Journal of international food & agribusiness marketing*, 5(1): 55-71.

Kofi T. (1973) 'A Framework for Comparing the Efficiency of Futures Markets.' *American Journal Of Agricultural Economics*, 55(4): 584.

Krichene N. (2008) 'Crude oil prices: Trends and forecast'. In: *IMF Working Paper No. 08/133*. Washington, DC: International Monetary Fund.

Kristoufek L. & Vosvrda M. (2014) 'Commodity futures and market efficiency'. *Energy Economics*, 42: 50-57.

- Kuhn, B. A. (1988) 'A note: Do futures prices always reflect the cheapest deliverable grade of the commodity?'. *The Journal of Futures Markets*, 8(1): 99.
- Lee, J. (2015) 'The Predictability of Futures Prices: A Comprehensive Study.' *Journal Of Accounting & Finance* (2158-3625), 15(7): 53-63.
- Levine D., Stephan D., Krehbiel M., & Berenson T. (2010) *Statistics for managers using Microsoft Excel* (6th edn.). Upper Saddle River, NJ: Pearson/Prentice Hall.
- Lucas, Robert E. B (1974) 'Hedonic Price Functions'. *Economic Inquiry*, 13: 157-178.
- Mitra S., & Boussard J. (2012) 'A simple model of endogenous agricultural commodity price fluctuations with storage.' *Agricultural economics*, 43(1): 1-15.
- Mohan S. & Love J. (2004) 'Coffee futures: role in reducing coffee producers' price risk.' *Journal Of International Development*, 16(7): 983-1002.
- Olanrele, O., Said, R. and Daud, M. (2014) 'Benchmark for REIT Performance in Malaysia Using Hedonic Regression Model.' *International Journal of Economics and Finance*, 6(9).
- Orge Pinheiro, C. & de Senna, V. (2016) 'Price Forecasting Through Multivariate Spectral Analysis: Evidence for Commodities of BM&Fbovespa'. *Brazilian Business Review (English Edition)*, 13(5): 129-157. Retrieved from Business Source Complete, EBSCOhost [Accessed 5 January 2017].
- Otero, Jesús and Costas Milas (2001) 'Modelling The Spot Prices Of Various Coffee Type.' *Economic Modelling* 18(4): 625-641.
- Paschke R., & Prokopczuk M. (2010) 'Commodity derivatives valuation with autoregressive and moving average components in the price dynamics.' *Journal of Banking and Finance*, 34(11): 2742–2752.
- Peterson S. & Flanagan A. (2009) 'Neural Network Hedonic Pricing Models in Mass Real Estate Appraisal'. *Journal Of Real Estate Research*, 31(2): 147-164.
- Rauch J., Krayzler M., Brunner B., & Zagst R. (2013) 'Pricing of derivatives on commodity indices.' *International Review Of Financial Analysis*, 29: 143-151.
- Schwager J. (1995) *Futures: Fundamental Analysis*. USA: John Wiley & Sons.
- Schwartz E. (1997) 'The stochastic behavior of commodity prices: Implications for valuation and hedging.' *Journal of Finance*, 52(3): 923–973.
- Susanti, Y. (2001) *The effect of exchange rate on Indonesian agricultural exports*. Doctoral dissertation. Oklahoma, USA: Oklahoma State University [published].

- Tomek W. (2000) 'Commodity prices revisited.' *Agriculture Resource Economics Review*, 29(2): 125–137.
- Tse Yiuman (1999) 'Price Discovery and Volatility Spillovers in the DJIA Index and futures markets.' *Journal of Futures Market*, 19: 911–930.
- Wolfers J. & Zitzewitz E. (2004) 'Prediction markets'. *Journal of Economic Perspectives*, 18: 107-126.
- Wooldridge J. (2012) *Introductory Econometrics: A modern approach* (5th edition). USA: Cengage Learning.
- Working H. (1962) 'New concepts concerning futures markets and prices.' *American Economic Review*, 52: 431–459.
- Yin K. (2001) *Case study research: Design and methods*. Thousand Oaks, CA: Sage.
- Zapata, H. and Fortenbery, T. (1996) 'Stochastic Interest Rates and Price Discovery in Selected Commodity Markets.' *Applied Economic Perspectives and Policy*, 18(4): 643-654.